

Received February 2, 2022, accepted February 17, 2022, date of publication February 22, 2022, date of current version April 5, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3153306

Improving Effectiveness of Different Deep Transfer Learning-Based Models for Detecting Brain Tumors From MR Images

SOHAIB ASIF^{1,2}, WENHUI YI¹, QURRAT UL AIN³, JIN HOU⁴, TAO YI⁵, AND JINHAI SI¹

¹Faculty of Electronic and Information Engineering, School of Electronic Science and Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

²School of Computer Science and Engineering, Central South University, Changsha, Hunan 410083, China

³School of Public Health, Central South University, Changsha, Hunan 410007, China

⁴Department of Pharmacology, School of Basic Medical Sciences, Xi'an Medical University, Xi'an, Shaanxi 710021, China

⁵School of Computer Science and Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

Corresponding author: Wenhui Yi (yiwenhui@mail.xjtu.edu.cn)


This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFA0706400 and Grant 2019YFA0706402, in part by the Pre-Research Funds for Equipment of China under Grant 61409220115, in part by the Natural Science Foundation of China under Grant 61275179 and Grant 60970815, in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Program 2016JM8047, in part by the Fundamental Funds for the Central Universities under Grant xjj2015112, in part by the Shaanxi Province Administration of Traditional Chinese Medicine under Grant 15-ZY036, and in part by the Xi'an Medical University State Fund Cultivation Project under Grant 2016GJF01.

ABSTRACT Early classification of brain tumors from magnetic resonance imaging (MRI) plays an important role in the diagnosis of such diseases. There are many diagnostic imaging methods used to identify tumors in the brain. MRI is commonly used for such tasks because of its unmatched image quality. The relevance of artificial intelligence (AI) in the form of deep learning (DL) has revolutionized new methods of automated medical image diagnosis. This study aimed to develop a robust and efficient method based on transfer learning technique for classifying brain tumors using MRI. In this article, the popular deep learning architectures are utilized to develop brain tumor diagnostic system. The pre-trained models such as Xception, NasNet Large, DenseNet121 and InceptionResNetV2 are used to extract the deep features from brain MRI. The experiment was performed using two benchmark datasets that are openly accessible from the web. Images from the dataset were first cropped, preprocessed, and augmented for accurate and fast training. Deep transfer learning models are trained and tested on a brain MRI dataset using three different optimization algorithms (ADAM, SGD, and RMSprop). The performance of the transfer learning models is evaluated using performance metrics such as accuracy, sensitivity, precision, specificity and F1-score. From the experimental results, our proposed CNN model based on the Xception architecture using ADAM optimizer is better than the other three proposed models. The Xception model achieved accuracy, sensitivity, precision specificity, and F1-score values of 99.67%, 99.68%, 99.68%, 99.66%, and 99.68% on the MRI-large dataset, and 91.94%, 96.55%, 87.50%, 87.88%, and 91.80% on the MRI-small dataset, respectively. The proposed method is superior to the existing literature, indicating that it can be used to quickly and accurately classify brain tumors.

INDEX TERMS Brain tumor classification, transfer learning, deep learning, magnetic resonance imaging, convolutional neural network.

I. INTRODUCTION

The brain is one of the most complex organs in the human body, controlling the entire nervous system and working with billions of cells [1]. The brain is the most sensitive organ

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo .

of our body. It controls core functions and is responsible for many regulatory functions of the human body such as memory, emotion, vision, and reaction. If certain tumors start to grow in the brain, these functions will be severely affected. This tumor is a primary brain tumor that resides in brain tissue, whereas the secondary tumor spreads through the bloodstream from other parts of the person's body into the

brain tissue [43]. A brain tumor is one of the worst diseases among other types of tumors due to its lower survival rate and aggressive nature. There are two types of brain tumors, malignant (cancerous) and benign (non-cancerous). Benign tumors do not contain cancer cells and grow slowly. It generally stays in one area of the brain. Malignant tumors are characterized by rapid spread to other brain tissues, making the patient's condition worse. The main symptoms of brain tumors include memory loss, frequent headaches, poor concentration, and problems with coordination. According to the National Brain Tumor Association, there are approximately 787,000 patients suffering from brain tumor diseases in the United States [2]. Patients reportedly have a survival rate of only 36%. By 2021, the estimated number of patients with brain tumors diagnosed is 84,170. Brain tumors are less common than other cancers such as breast and lung cancer. However, brain tumors are still the 10th most common cause of death in the world.

There are different ways to treat a brain tumor, depending on the size and type of tumor. Brain tumors vary in size they are often difficult to detect. In the early stages, it may not be possible to accurately measure the size and resolution of a brain tumor, making it even more difficult to detect. It is worth noting that early detection of brain tumors can improve the survival rate of patients. Diagnosis of brain tumors is very difficult compared to tumors in other parts of the human body. Since the brain is full of blood brain barriers, normal radioactivity indicators cannot capture tumor cell overactivity [3]. Recent imaging techniques have achieved great success in the field of medical imaging and can be used to diagnose dangerous human diseases, such as brain tumors [4], skin cancer [5], and stomach cancer [6]. For brain tumors, magnetic resonance imaging (MRI) and computer tomography (CT) scans are considered to be the best diagnostic systems for detecting brain tumors [44]. However, compared to CT images, MRI scans based on tumor texture and shape information are more useful. MRI is preferred because it is the only non-invasive medical imaging process that provides high-resolution images of brain tissue.

Computer-aided diagnosis (CAD) systems can help radiologists in the clinic for early detection of brain tumors. Currently, researchers have developed several automated systems for brain tumor detection, such as supervised and unsupervised machine learning techniques [7], transfer learning [8], [41], [42] and deep neural models [9]. The latest advances in artificial intelligence (AI) and deep learning (DL) have made great success in medicine. This allows the doctor to diagnose the disease early. Convolutional neural networks (CNNs) are widely used in various CAD systems [10], [11]. The CNN contains three basic layers, each convolution layer extracts features, a pooling layer is used to reduce the dimensions of the feature map, and a fully connected (FC) layer performs the classification. Deep learning-based methods automatically extract features, which can significantly improve performance. However, these methods require large amounts of data to improve accuracy, and obtaining such data is a difficult task in

the medical field. To solve this problem, it is necessary to develop an automatic CAD system for diagnosing life threatening diseases such as cancer, which is the main cause of death of patients worldwide. This paper proposes a new automated deep learning system for examining MRI of the brain and providing early diagnosis with improved performance. The key contributions of the presented study are as follows:

1. A novel deep learning-based system is proposed that uses state-of-the-art deep learning architectures such as Xception, NasNet Large, DenseNet121, and InceptionResNetV2 using MRI of brain tumors and applied transfer learning technique on the dataset.
2. The MR images have been improved during the preprocessing phase and various techniques such as data augmentation, three different optimizers (ADAM, SGD and RMSprop) and the L2 regularizer were used to improve classification performance.
3. The performance of the proposed system has been tested on two different well-known brain MRI datasets.
4. The proposed system has been compared with competing models in terms of various performance metrics, such as accuracy, precision, sensitivity, specificity, F1-score, Matthews Correlation Coefficient (MCC), and error rate

II. RELATED WORK

CNN has been widely used to solve different problems, but its performance is very good for image processing in health applications. In the past few years, various methods have been developed to identify brain tumors in MRI images based on DL. Most of them focus on binary classification for the detection of brain tumors. The author in [12] used brain tumor images for training, and created two main methods of recognition using CNN, and achieved the highest accuracy rate of 91.43%. Deepak and Ameer [10] used the Google Net architecture to classify brain MRI images. They achieved 98% classification accuracy using CNN, SVM, and KNN classifiers. Ahmet inar *et al.* [13] modified the Res-Net50 CNN model and compared its accuracy with other pretrained models such as GoogleNet and AlexNet. The modified ResNet 50 model achieved 97.2% classification accuracy to classify brain tumors. Sajjad *et al.* [14] expanded the data set and fine-tuned the proposed CNN method. The proposed method is applied to the original data set and the enhanced data set respectively. The enhanced data set achieves an accuracy of 94.58%. Kumar *et al.* [15] used a LSTM network and machine learning techniques to classify brain tumors. They used data augmentation to increase the dataset and support vector machines as a classifier. They achieved a classification accuracy of 78.33% in the brain tumor dataset. Shree and Kumar [16] used GLCM and a probabilistic neural network classifier for feature extraction and classification to classify MRI images of the brain as normal and neoplastic, and achieved 95% accuracy. Nooren *et al.* [17] used DenseNet 201 and Inception V3 to diagnose brain tumors. Their

approach developed linked multistage feature extraction of tumors and generated accurate predictions. The Inception V3 achieved 99.34% accuracy and the DenseNet 201 achieved 99.51% accuracy. Chelghoum *et al.* [18] used nine pre-trained deep learning models to classify BT with Transfer Learning (TL). They used fine-tuned TL-based models to experiment with MR data. The proposed model achieved an accuracy of 98.71% for MRI classification. Badža and Barjaktarović [19] proposed the CNN model. They use two convolutional layers and two fully connected layers for feature extraction and classification of brain tumors. Their accuracy of classifying brain tumors reached 97.39%. F. J. Díaz-Pernas *et al.* [20] proposed a multi-scale deep CNN that can analyze tumor MRI. They evaluate the performance of the proposed model on the MRI image dataset, and a classification accuracy of 97.3% was achieved. Rehman *et al.* [21] utilized three different pre-trained CNN models with transfer learning methods to classify brain tumors. The highest accuracy rate obtained by VGG16 is 98.67%. Taló *et al.* [22] applied a transfer learning method on a pretrained CNN ResNet34 model to classify normal and abnormal brain MRI images. They used a data augmentation technique to increase the images and achieved 100% accuracy. Ahmad *et al.* [23] proposed a combination of ANN model and Gray Wolf Optimizer (GWO) optimization techniques to classify brain tumor images as normal or abnormal. They obtained 98.91% classification accuracy with GWO-ANN. Saxena *et al.* [24] used the transfer learning method on the Inception V3, ResNet-50 and VGG-16 models to classify brain tumor images. The proposed model received a maximum accuracy of 95%. Kumar *et al.* [25] proposed the deep CNN network using ResNet-50 and trained it with brain MR images. The proposed model achieved 97.48% accuracy with data augmentation technique. Khwaldeh *et al.* [26] proposed a CNN model for classifying MRI images of the brain into normal and abnormal. They modified the AlexNet CNN model and got 91% accuracy. Abiwinanda *et al.* [27] used five different pre-trained and one sequential CNN model to classify brain tumor images and achieved the highest classification accuracy of 84.19%. Rehman *et al.* [4] proposed a three-dimensional CNN model for extracting features from MRI of the brain. They extracted features using CNN and then ANN is used for classification. For three different data sets, the accuracy of the proposed method is 92.67%, 96.97%, and 98.32%, respectively. Díaz-Pernas *et al.* [28] proposed a multi-path CNN model for the segmentation of a brain tumor. They evaluated the performance of the proposed model on a publicly available MRI dataset and obtained an accuracy of 97.3%. Cheng *et al.* [36] performed segmentation of the dataset and used KNN and SVM classifier and achieved 91.28%. Afshar *et al.* [37] proposed a Capsnet CNN model and achieved 90.89% accuracy on brain MRI dataset. Soltaninejad *et al.* [38] used random forest classifier on brain MRI dataset and achieved 86% accuracy. Mehrotra *et al.* [39] proposed AlexNet model and performed transfer learning on the dataset and achieved 99.04% accuracy. Kang *et al.* [40]

combined three different CNN models and built an ensemble model, achieving 98.83% accuracy.

III. PROPOSED METHODOLOGY

Early diagnosis of brain tumors is of great significance to clinical diagnosis and effective treatment. Manual brain tumor detection is a complex task that depends on expertise in identifying brain tumors. In this work, an effective deep learning-based framework is proposed to automatically classify brain tumors with minimal doctor intervention. The purpose of this study is to use DL algorithms and TL techniques to improve the accuracy of MR image identification in the brain. The workflow of our proposed brain tumor classification method is shown in Fig. 1. The proposed framework model includes four stages. First, the input MR image is preprocessed (brain cropping and resizing, data splitting and normalization). Second, the data augmentation technique is used to increase the size of the dataset. Third, we investigated the four unique DL models, such as Xception, NasNet Large, DenseNet121, and InceptionResNetV2, using BT's preprocessed MR images and applied TL technique to extract features. The features extracted by the CNN models are classified using the softmax layer.

A. DATASET DESCRIPTION

We conducted a set of experiments on two different publicly available brain MRI data sets. BR35H: Brain Tumor Detection 2020 (BR35H) [29] is the first brain MRI dataset downloaded from the Kaggle website. For experimental work, we named this dataset MRI-large. The MRI-large dataset contains 1500 images containing the tumor, and the remaining 1500 images are normal. Samples belonging to the normal and tumor class in the dataset are shown in Fig. 2.

The second dataset consists of open access brain tumor MRI [30], and for experimental work, we call this dataset MRI-small, which includes two classes: tumor and normal. The MRI-small data set contains 253 images from 253 patients. The dataset includes 155 tumor images and 98 normal images. We balanced the sample size across the class by increasing the number of MRIs of normal brain tumors to use the dataset more efficiently. The number of normal class images was increased from 98 to 155 using the image augmentation method. Fig. 3 shows examples of MRI images from the MRI-small dataset.

B. DATA PREPROCESSING

Data preprocessing is the most important factor in image analysis. Almost all images in the brain's MRI dataset contain unnecessary space and areas, noise, and missing values. This can reduce the performance of the classifier. Therefore, it is necessary to remove unwanted areas and noise present in the MRI image. We use the cropping method to crop the image by calculating extreme points and finding contours [31]. Fig. 4 shows the process of cropping an image using extreme point calculation. First, we load raw images from the brain tumor MRI dataset for preprocessing. After

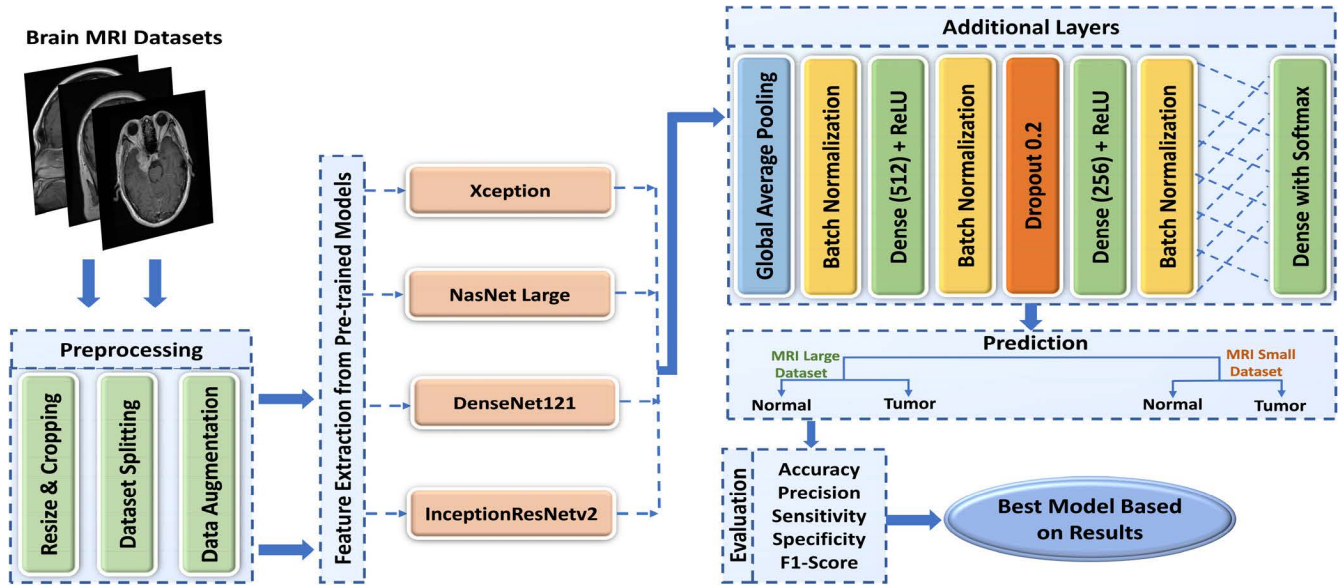


FIGURE 1. Workflow of the proposed method for predicting MRI images of normal and tumor.

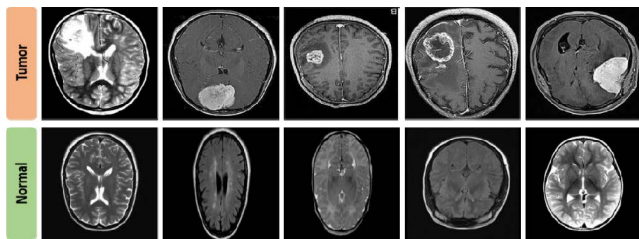


FIGURE 2. Examples of MR images from the MRI-large dataset. The first line represents the MRI diagnosis of tumor, while the second line represents normal.

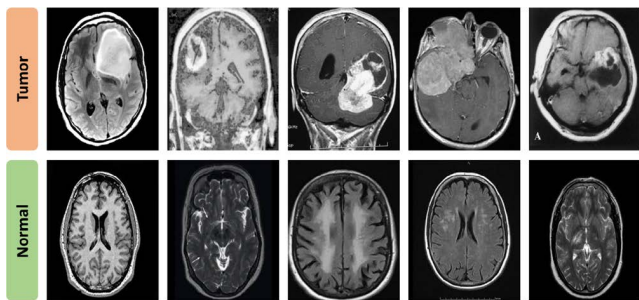


FIGURE 3. Examples of MR images from the MRI-small dataset.

that, all RGB images are converted to grayscale. After that, threshold processing is applied to convert the gray-scale image into binary. In addition, we remove small areas of noise using dilation and erosion operations. After that, we find the contour in the threshold image, and then grab the largest contour. Then we find the extreme points based on the largest contour. Finally, crop the image according to the contour and extreme points. Fig. 5 shows a sample of images from the dataset after the cropping process.

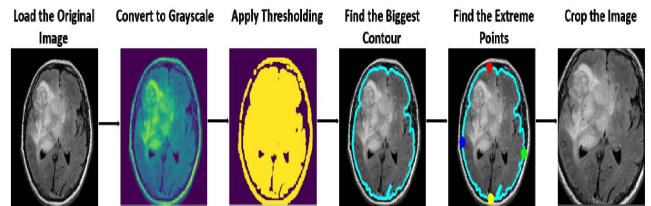


FIGURE 4. Image cropping process.

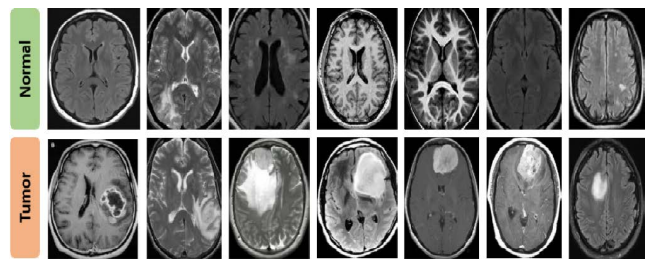


FIGURE 5. Sample of images from the MRI dataset after the cropping process.

The MR images in the dataset have different sizes, so it is recommended to adjust them to the same height and width for best results. In order to allow all architectures used in this study to accept a common size, we initially adjusted all brain tumor images to a (224 x 224 x 3) shape. Both datasets used in the experiment are divided into training, validation, and testing. Table 1 shows the total number of images from the two datasets for each category used for training, testing, and validation, and Fig. 6 shows the number of images for each category (normal and tumor). Data normalization is also used to convert the input image to a range of pixel values with an interval of [0,1].

TABLE 1. The number of images used in the training, validation and testing stages without data augmentation.

Dataset	Training Set	Validation Set	Test Set	Total
MRI-large	1923	475	602	3000
MRI-small	198	50	62	253

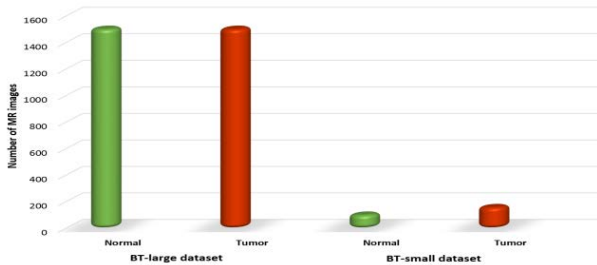


FIGURE 6. The number of MRI images per class in the dataset.

TABLE 2. The number of images with and without data augmentation.

Dataset	Training Set without augmentation	Training Set with augmentation
MRI-large	1923	7692
MRI-small	198	792

C. DATA AUGMENTATION

Due to the small number of images in the brain MRI dataset, we used image augmentation technique. It is also used to solve the problem of model overfitting. Due to the deep nature of DL models, if the size of the data set is small, there is a high risk of overfitting. Additional images were created using data augmentation to overcome this drawback. In this work, we used three augmentation strategies to generate a new training set (rotation, horizontal flip and translation). According to reports, data augmentation can improve the classification accuracy of the DL algorithms instead of collecting new data. It is worth noting that three augmentation strategies were used in order to generate a new training set. Initially, 1923 images from the MRI large dataset and 198 images from the MRI small dataset were assigned to train the model, yielding 7692 training images from the MRI large dataset and 792 images from the MRI small dataset after data augmentation. This is 4 times larger than the original training images. Table 2 summarizes the number of training images after data augmentation.

D. DEEP FEATURE EXTRACTION USING PRE-TRAINED CNN

1) CNN

As mentioned earlier, CNNs are popular because of their improved image classification performance. CNN automatically collects features from the input data. It is one of the well-known DL architectures in which each layer is connected in a feed-forward manner. Deep architecture

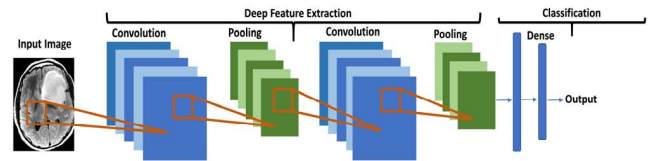


FIGURE 7. CNN architecture.



FIGURE 8. Concept of transfer learning.

helps these networks learn various and complex functions that a simple neural network cannot learn. CNN is the core of computer vision, and it has many applications, including object classification, surveillance, and medical imaging. Compared to other neural classifiers, it contains an internal filter, so it uses a very simple and relatively small preprocessing. A typical CNN architecture consists of the following layers: (i) Convolution Layer (ii) Pooling Layer (iii) Activation function (iv) Dense Layer to classify the input image. The architecture overview of CNN is shown in the Fig. 7.

2) TRANSFER LEARNING

TL is a deep learning technique that uses a model pre-trained on a large dataset for a problem as an initialization for a model trained on a different dataset. CNNs tend to perform better with larger datasets than smaller ones. TL can be useful in CNN applications where the dataset is small. Recently, TL has been used for object detection, medical imaging and image classification. Fig. 8 shows the concept of transfer learning. Models trained on large datasets such as ImageNet can be used as feature extractors for a variety of applications using smaller datasets such as brain MRI datasets. The benefits of TL are faster training processes, prevention of overfitting, training with less data, and improved performance. The pre-trained CNN models used in our study are Xception, NasNet Large, DenseNet121, and InceptionResNetV2.

E. PROPOSED DEEP TRANSFER LEARNING MODELS FOR CLASSIFICATION

The main goal behind the development of our proposed model is to automatically distinguish people with brain tumors, while reducing the time required for classification and improving accuracy. We propose a novel and robust DL framework for detecting brain tumors using MRI datasets. Due to the lack of data composed of brain MRI, it is

TABLE 3. Characteristics of the deep CNN architectures with top 5 error rate considered for this work.

Model	Input Size	Custom Input Size	Parameters (M)	Size (MB)
Xception	299×299	224×224	22.9	88
NasNet Large	331×331	224×224	88	343
DenseNet121	224×224	224×224	8	33
InceptionResNetV2	299×299	224×224	56	215

usually not feasible to train the model from scratch using randomly initialized weights. Therefore, to avoid these problems, we use the TL technique because our MRI dataset is not very large. For the source task, we use four network architectures, namely: Xception, NasNet Large, DenseNet121, and InceptionResNetV2. We extract the deep features of the brain MRI dataset using the fixed weights of each model pre-trained on the large ImageNet dataset. The pre-trained models mentioned above are initialized with their pre-trained weights on the ImageNet dataset. After the feature extraction block, the extracted deep features were transferred to 3 dense or FC layers using the global average pooling layer to average the spatial dimensions of the features. Global average pooling is mainly used to replace the fully connected layer in CNN, which helps to minimize over-fitting. The first dense layer contains 512 neurons followed by the ReLU function, the second dense layer contains 256 neurons followed by the ReLU function, and the last dense layer is followed by a softmax activation function to classify the image. The batch normalization layer is used after each FC layer to normalize the extracted features to mean and standard deviation. The first dense layer is equipped with L1 and L2 regularizers. In addition to regularization, a dropout method and early stopping criterion was also applied to prevent model overfitting. The models are trained for 50 epochs with a batch size set to 64. Each model was trained with three different optimizers: Adam, Stochastic Gradient Descent (SGD), and RMSprop. Table 3 summarizes the important characteristics and some key features of the adopted deep CNN models. The detailed explanation of the model used in this study is mentioned in the following section.

1) XCEPTION

Xception is one of the most advanced deep learning model architectures based on a depth-separable convolution layer developed by Chollet [32], also known as the successor to the Inception network. The network consists of 36 convolutional layers, excluding the last fully connected layer, and can be divided into 14 different modules. Each module has linear residual links around them, with the exception of the start and end modules. This architecture is a stack of depth-separable convolution layers consisting of depth-wise convolutions - spatial convolution is performed on each

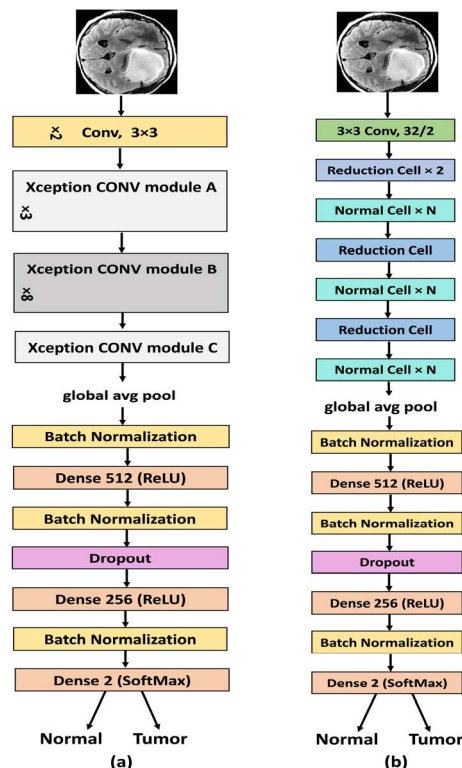


FIGURE 9. (a) Basic architecture and customization in Xception model for brain tumor classification (b) Basic architecture and customization in NasNet Large model for brain tumor classification.

input channel to map spatial relationships. Then a 1×1 depth convolution is performed to capture the cross-channel correlation. After performing this operation, the correlation can be considered as a mapping of $2D + 1D$ instead of a $3D$ mapping. In extreme versions of the inception module, depth and pointwise convolutions are followed by ReLU non-linearity. Xception is superior to inceptionV3, VGG and ResNet in the classification of the ImageNet dataset. Fig. 9 (a) highlights the basic architecture of the Xception model and its customization, which was finally deployed in this work to obtain classification results.

2) NASNET LARGE

NASNet is an architecture created using neural architecture search algorithms [33]. This idea is realized by the concept of NAS developed by the Google ML team. Their method is based on reinforcement learning. In this network, the efficiency of the child blocks is checked by the parent block and the architecture of the neural network is tuned. Several changes have been made based on optimizer functions, weights, regularization methods, etc. to improve network efficiency. The network components include a Recurrent Controller Neural Network (CRNN) and a CNN block. A block is the smallest unit of NASNet architecture, and a cell is a combination of blocks. The network search space is created by dividing the network into cells and further dividing into blocks. Possible operations for blocking include regular

convolution, separable convolution, pooling, and identity mapping. In this work, NasNet was chosen to identify brain tumor patients and normal patients because the network has a scalable architecture for image classification and is composed of normal cells and reduced cells. Fig. 9 (b) shows the basic architecture of the NasNet Large model and its settings, which were finally deployed in this work to obtain classification results.

3) DENSENET121

The Dense Convolution Network (DenseNet) is a pre-trained deep learning model that uses feedforward to connect each layer to all subsequent layers [34]. Traditional L-layer CNNs have L connections, DenseNet has $(L \times (L+1))/2$ direct connections. Each layer in the model contains a feature map. The feature map of each layer is used as the input of the next layer. By connecting all layers directly to each other, it provides maximum information transfer within the network. The main advantages of DenseNet are that it significantly reduces the number of parameters, reduces gradient runaway, enhances feature diffusion, and promotes feature reuse. Compared with traditional CNN. DenseNet requires fewer parameters because the feature map is not learned redundantly. In addition, DenseNet reduces the chance of overfitting by applying regularization. DenseNet121 contains four dense blocks, and each dense block contains 6, 12, 24, and 16 convolution blocks. Fig. 10 (a) highlights the modified architecture of the Densenet121 model, which was finally deployed in this work to obtain classification results.

4) INCEPTIONRESNETV2

The InceptionResNetV2 architecture is based on the Inception block and uses transformation and merging functions for feature extraction [35]. It provides the highest performance with less computational cost compared to inceptionResNetV1. It is based on a combination of residual learning and inception block. Convolution filters of multiple sizes are combined by a residual connection. Residual connections not only avoid the problem of degradation, but also shorten the training time. This network uses three types of blocks: Stem block, InceptionResNet block, and Reduction block to significantly improve recognition performance. It is a deep network that connects one main block, 5 Inception ResNetA, 10 Inception ResNet-B, 5 Inception ResNet-C, one Reduction-A block, and one Reduction-B block. Fig. 10 (b) shows a modified architecture of the InceptionResNetV2 model for classifying brain MRI.

F. REGULARIZATION AND OPTIMIZATION TECHNIQUES

The main goal of this work is to create a better model for classifying patients with brain tumors. In order to avoid overfitting during training, a regularization function (L2) is used. This is a technique to avoid overfitting the algorithm and avoid overfitting the coefficients so that they fit perfectly as the model becomes more complex. In addition

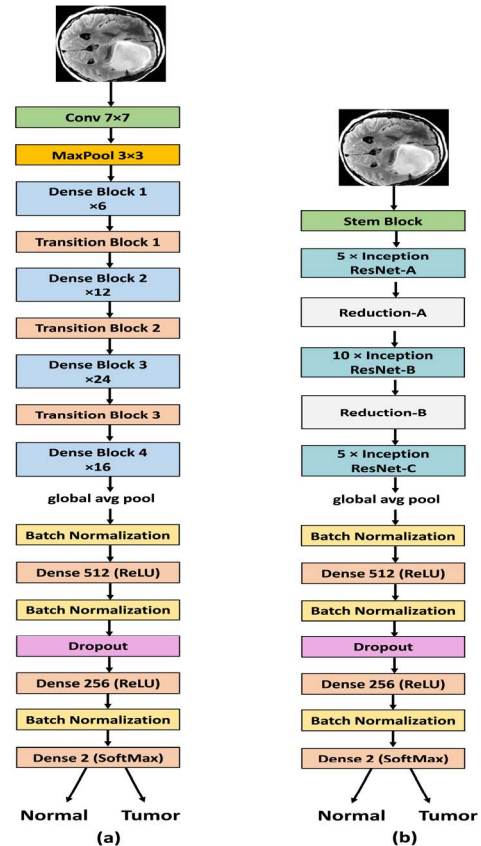


FIGURE 10. (a) Basic architecture and customization in DenseNet121 (b) Basic architecture and customization in InceptionResNetV2.

to regularization, batch normalization and global average pooling are also applied to prevent the model from overfitting. In the preprocessing and training stages, many methods have been used to prevent overfitting. Initially, data augmentation techniques are used to improve the performance of the model. Early stopping technique is used to stop the model if no updates occur to the validation accuracy and avoid overfitting the system. It provides the possibility to stop the iteration where the deviation of the validation error begins. Finally, dropouts are used not only to speed up the learning of the algorithm, but also to significantly reduce overfitting. The cross-entropy loss function was used to minimize the loss. It is used to optimize the parameter values used in our models. Three optimization algorithms Adam, SGD and RMSprop are used to train the model. These optimizers are used to identify tumors from MRI and are compared to identify the best optimizers for detecting brain tumor disease.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

This study was conducted to diagnose patients with brain tumor symptoms with the help of brain MRI. Various deep learning models (Xception, NasNet Large, DenseNet121, and InceptionResNetV2) were trained and tested with multiple optimizers. Models are trained multiple times using a variety of well-established optimizers such as Adam, SGD, and RMSprop to get the best possible trained system.

TABLE 4. Summary of Hyperparameters used for training the models.

Performance Measure	Xception	NasNet Large	DenseNet 121	InceptionResNetV2
No. of epochs	50	50	50	50
Batch Size	64	64	64	64
Image Size	224×224	224×224	224×224	224×224
Optimizers	Adam, SGD, RMSprop	Adam, SGD, RMSprop	Adam, SGD, RMSprop	Adam, SGD, RMSprop
Activation Function	Softmax	Softmax	Softmax	Softmax
Learning Rate	1e-4	1e-4	1e-4	1e-4
Decay Rate	1e-5	1e-5	1e-5	1e-5
Dropout Rate	0.2	0.2	0.2	0.2
Regularizer	1e-4	1e-4	1e-4	1e-4

A. EXPERIMENTAL SETUP

To implement the proposed models and obtain results, we used the Python 3.7 programming language, Keras 2.3.1 and TensorFlow 2.0 libraries. The matplotlib and seaborn libraries were used for visualization. System specifications: Intel(R) Core (TM) i5 @ 2.50 GHz, 12GB RAM, NVIDIA Tesla K80 GPU and window 7 installed. For statistical results, the images of the input classes were augmented using the Keras Augmentor API.

B. HYPERPARAMETERS TUNING

The main goal of this task is to design the optimal model for the classification of brain MRI. The set of parameters that can affect the training of the model and give the best results is called hyperparameters. These parameters include number of epochs, dropout number, activation function, batch size, learning rate, etc. After several trials during the experiment, we fixed the learning rate, batch size, and regularization factor. The performance of the proposed brain tumor detection is done using different pre-trained models, such as Xception, NasNet Large, DenseNet121 and InceptionResNetV2. Each model has been evaluated for 50 epochs using various optimizers. Table 4 shows the parameters used to train the models.

C. PERFORMANCE EVALUATION METRICS

This section describes the metrics used to quantify the classification performance of the models. There are several ways to evaluate the performance of a classifier, but we used a confusion matrix-based metric to validate the results. The performance metrics such as accuracy, precision, sensitivity, specificity, F1-score, Matthews Correlation Coefficient (MCC), NPV and error rate are used to evaluate the predictive strength of the model [45], [46].

Accuracy: Accuracy is the ratio of correctly classified images to the total number of images.

Sensitivity: It is used to correctly identify patients suffering from a certain disease.

Specificity: It is used to correctly identify people without the disease

Precision: It shows the model reliability is classifying the images as positive.

F1-score: It combines the precision and recall by taking harmonic mean.

MCC: It is a more reliable statistical metric for binary classification problems.

NPV: It calculates the percentage of true negative predictions that do not have the disease.

FPR: The ratio of false positive and total negative.

FNR: The ratio of false negative and total positive.

$$\text{Accuracy (ACC)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Sensitivity (SEN)} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Precision (PRE)} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Specificity (SPE)} = \frac{TN}{TN + FP} \quad (4)$$

$$\text{F1-Score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (5)$$

$$\text{NPV} = \frac{TN}{FN + TN} \quad (6)$$

$$\text{MCC} = \frac{((TP \times TN) - (FP \times FN))}{\sqrt{((TP + FP)(TP + FN)(TN + FP)(TN + FN))}} \quad (7)$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP + TN} \quad (8)$$

$$\text{False Negative Rate (FNR)} = \frac{FN}{FN + TP} \quad (9)$$

To assess these indicators, we needed to calculate the following values - true positive, false negative, true negative and false positive.

TP: Number of images correctly classified as brain tumor patients.

FN: Number of images misclassified as healthy.

FP: Number of images misclassified as brain tumor patients.

TN: Number of images correctly classified as healthy.

D. RESULTS AND ANALYSIS

This study was conducted to diagnose brain tumor patients using MRI of the brain. The empirical results are obtained for two different datasets (MRI-small and MRI-large) for brain tumor classification tasks. This article uses a TL-based deep learning models to accurately identify patients as normal and tumor. The system is trained using various deep learning pre-trained networks such as Xception, NasNet Large, DenseNet121, InceptionResNetV2 to access the highest accuracy of the system. The system is trained multiple times against all of the above pre-trained networks using a variety of acclaimed optimizers such as ADAM, SGD, and RMSprop to classify the two brain tumor types from MR images. The reason for performing this study with a large number of parameters is to find the best combination of model and

TABLE 5. Performance of the Xception model of each optimizer using various evaluation metrics.

Optimizers	ACC	SEN	PRE	SPE	F1-Score	NPV	MCC	FPR	FNR
ADAM	99.67	99.68	99.68	99.66	99.68	99.65	99.33	0.0034	0.0032
SGD	96.35	98.39	94.74	94.16	96.53	98.21	92.75	0.581	0.161
RMSprop	99.34	99.04	99.68	99.66	99.35	98.98	98.67	0.0034	0.0096

TABLE 6. Comparative analysis of Adam, RMSprop and SGD optimizers used in NasNet Large models.

Optimizers	ACC	SEN	PRE	SPE	F1-Score	NPV	MCC	FPR	FNR
ADAM	99.34	99.36	99.36	99.31	99.36	99.31	98.67	0.0069	0.0064
SGD	98.34	98.71	98.08	97.94	98.40	98.62	96.68	0.0206	0.0129
RMSprop	99.00	98.71	99.35	99.31	99.03	98.63	98.01	0.0069	0.0129

TABLE 7. Performance of the DenseNet121 model of each optimizer in terms of different evaluation metrics.

Optimizers	ACC	SEN	PRE	SPE	F1-Score	NPV	MCC	FPR	FNR
ADAM	99.00	99.36	98.72	98.63	99.04	99.31	98.01	0.0137	0.0064
SGD	94.68	95.82	94.01	93.47	94.90	95.44	89.37	0.0653	0.0413
RMSprop	98.84	99.68	98.10	97.94	98.88	99.65	97.68	0.0206	0.0032

optimizer for the input data. We report quantitative results along with confusion matrices for each adopted network architecture.

1) EXPERIMENTATION 1: CLASSIFICATION ACCURACY ON MRI-LARGE DATASET

In our research, a total of four models were developed, and the performance of each model was evaluated based on the measures discussed in Section 4.3. we present and discuss the results of brain tumor detection on the considered MRI dataset using our TL models with three different optimizers, i.e. Adam, SGDM and RMSprop. To classify two brain tumor types from MR images. First, the Xception transfer learning model was tested for each optimizer. The detailed classification results obtained using the Xception model are compared in terms of various indicators and are summarized in Table 5. As shown in Table 5, the best classification performance is achieved using ADAM. Brain MRI scans are classified with an accuracy of 99.67%. Then, using SGD and RMSprop, the MRI was classified with 96.35% and 99.34% accuracy, respectively. We can see that the Xception model adapted all three optimizers very well and synthesized the highest accuracy of 99.67% with the ADAM optimizer. These results indicate the superiority of the Xception model in brain tumor classification.

The NasNet large architecture was used as the second method of brain tumor classification using three optimizers. The results related to NasNet large are shown in Table 6. Here, we have achieved the best performance of 99.34% using the ADAM optimizer. Brain tumors were classified using other optimizers SGD and RMSprop with very high performance of 98.34% and 99.00%, respectively.

As mentioned earlier, the Xception and NasNet Large models are 88MB and 343MB in size, respectively. On the

other hand, the model size of DenseNet121 is 33MB, which is actually much smaller than other models. The DenseNet121 architecture requires fewer parameters than a traditional CNN. The connection pattern eliminates the need to learn redundant features maps. Xception and NasNet large have approximately 22.9M parameters and 88M parameters, while the DenseNet121 network has approximately 8M parameters. DenseNet121 model was tested to classify brain tumors. As with the others, three different optimizers were used to test the performance of the model. Table 7 shows the obtained classification performance. The best achievement in classification was achieved using ADAM. We can say that the classification performance obtained using other optimizers are also quite good.

Finally, the InceptionResNetV2 transfer learning model was tested for each optimizer. Table 8 shows the obtained classification performance. As shown in Table 8, the best classification performance is achieved using ADAM. Brain MRI are classified with an accuracy of 99.67%. Moreover, RMSprop was second with an average accuracy of 99.50%. For the SGD Optimizer, InceptionResNetV2 showed surprisingly bad results with an accuracy of 87.38%. This is worse than all other scenarios in this study.

Fig. 11 shows an accuracy comparison of different TL models using different optimizers. In general, almost every model is well adapted to the ADAM optimizer. We can quickly see that all modified TL models with the ADAM optimizer are probably the best combination of accuracy and other metrics for evaluating performance.

The detailed classification results obtained from all TL models using the ADAM Optimizer are compared in terms of various metrics and summarized in Table 9. All values are given as percentages and the best results are shown in bold. It can be seen that the Xception model achieved

TABLE 8. Performance of the InceptionResNetV2 model of each optimizer in terms of different evaluation metrics.

Optimizers	ACC	SEN	PRE	SPE	F1-Score	NPV	MCC	FPR	FNR
ADAM	99.67	99.36	1.000	1.000	99.68	99.32	99.34	0.0000	0.0064
SGD	87.38	79.10	95.72	96.22	86.62	81.16	76.10	0.0378	0.2090
RMSprop	99.50	99.68	99.36	99.31	99.52	99.66	99.00	0.0069	0.0032

TABLE 9. Performance comparison of different deep TL models for detecting brain tumors using different metrics.

Model	ACC	SEN	PRE	SPE	F1-Score	NPV	MCC	FPR	FNR
Xception	99.67	99.68	99.68	99.66	99.68	99.65	99.33	0.003	0.003
NasNet Large	99.34	99.36	99.36	99.31	99.36	99.31	98.67	0.006	0.006
DenseNet121	99.00	99.36	98.72	98.63	99.04	99.31	98.01	0.013	0.006
InceptionResNetV2	99.67	99.36	1.000	1.000	99.68	99.32	99.34	0.000	0.006

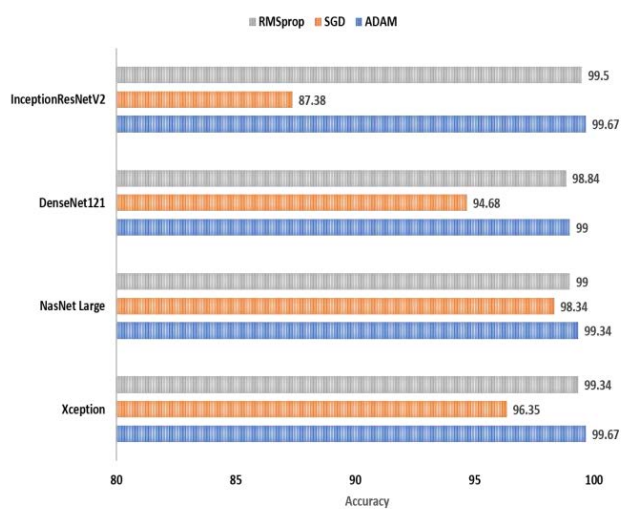


FIGURE 11. Comparison of the accuracy of different TL models with different activation functions.

the highest performance with 99.68% precision, 99.66% specificity, 99.68% F1-score, 99.67% accuracy. It should also be noted that the obtained sensitivity is significantly higher (i.e. 99.68%). The InceptionResNetV2 model was found to be the second-best performance for brain tumor prediction, achieving 100% precision, 99.36% sensitivity, 100% specificity, 99.68% F1-score and 99.67% accuracy. NasNet Large performs reasonably well, reaching 99.34% accuracy on the test set, and the precision, sensitivity, specificity, and F1-score are 99.36%, 99.36%, 99.31%, and 99.36%, respectively. DenseNet121 achieves 99% accuracy on the test set, DenseNet121 achieves precision, sensitivity and F1-scores of 98.72%, 99.36% and 99.04%. Based on the results of the experiment, we can draw the following conclusions: Deep TL methods are very effective for classifying brain MRI. The Xception and InceptionResNetV2 models perform classification with similar accuracy rates. Xception is slightly more successful than InceptionResNetV2. Xception and InceptionResNetV2 models have achieved almost the same results in all evaluation indicators. According to observations,

TABLE 10. Class-wise Precision, Sensitivity and F1-Score for all the models.

Model	Class	PRE	SEN	F1-score
Xception	Tumor	99.6	99.6	99.6
	Normal	99.6	99.6	99.6
NasNet Large	Tumor	99.3	99.3	99.3
	Normal	99.3	99.3	99.3
DenseNet121	Tumor	98.7	99.3	99.0
	Normal	99.3	98.6	98.9
InceptionResNetV2	Tumor	100	99.3	99.6
	Normal	99.3	100	99.6

Xception and InceptionResNetV2 are superior to other TL models in almost all performance indicators (including accuracy, sensitivity, and specificity). In this case, we can see that these two models achieve the same classification accuracy rate of 99.67% on the test data set, which is used to classify brain tumors and healthy patients. Both models show high sensitivity (99.68% and 99.36%, respectively) and specificity (99.66% and 100%, respectively), which are two very important performance indicators in medical applications. It can be seen that the Xception model gives very good results in brain tumor classification. In addition, the results obtained using the InceptionResNetV2 model are close to the results of Xception, which is quite good.

Class-wise performance of TL models are presented in Table 10. The various classes used in this study are tumor and normal. From the table it can be concluded that the models showed good results in the tumor and normal class. The Xception model also achieved the highest sensitivity score of 99.6% in the “tumor” class compared to all other models. InceptionResNetV2 achieves 100%, 99.3% and 99.6% precision, sensitivity and F1-score for tumor class.

The analysis of the proposed TL models is presented using loss and accuracy curves. Fig. 12 and Fig. 13 shows the training performance in terms of training loss, training accuracy, validation loss, and validation accuracy obtained by four different TL models using ADAM optimizer at different

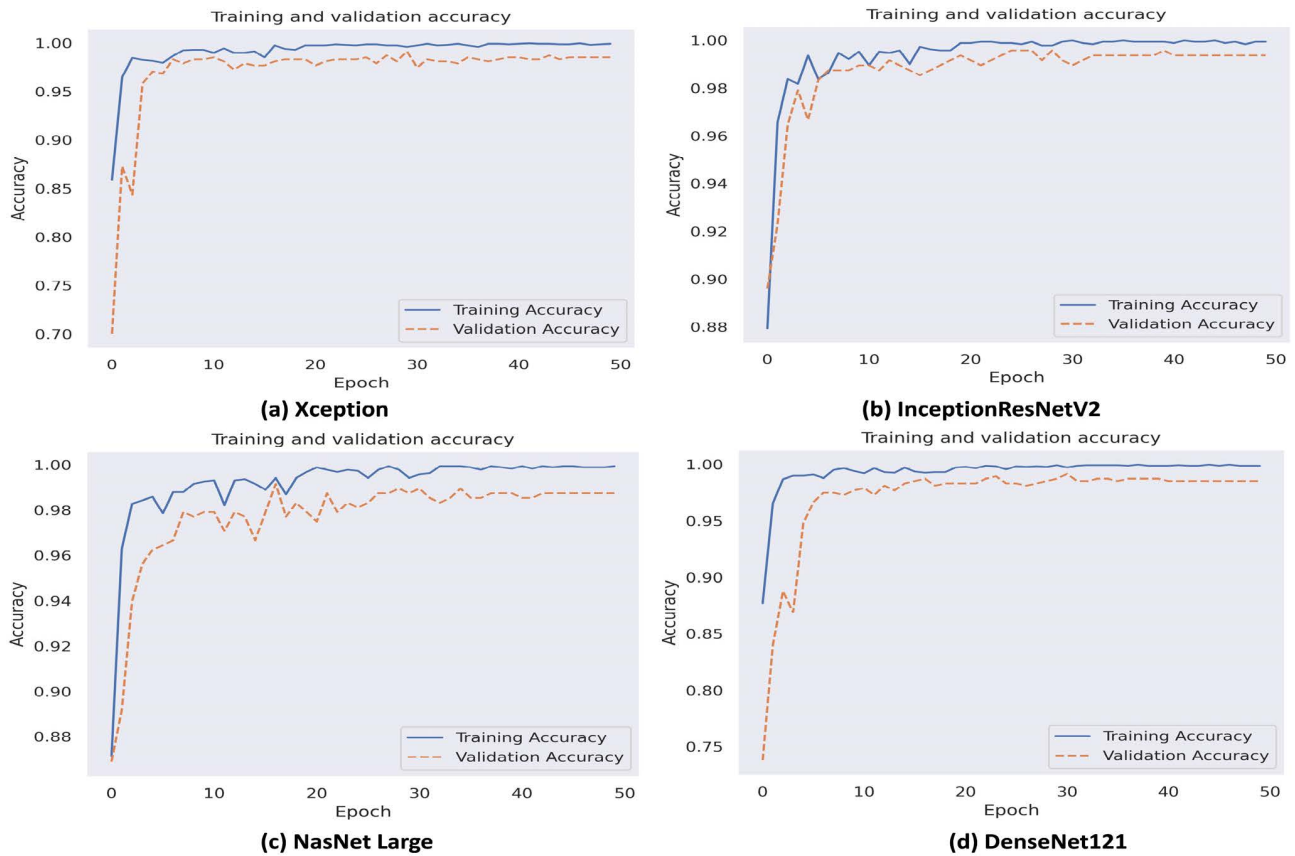


FIGURE 12. Epochs versus accuracy graph for all the models.

epochs. The models converge well and reaches the highest accuracy with minimal training and validation losses.

In contrast, the accuracy of the model increases with the number of epochs. The learning curves also show that the models are not overfitting to the training dataset. This means that at each epoch, the model is learning the given input very well. This is primarily due to the use of dropout regularization techniques applied to the proposed TL models and image augmentation to address the shortage of available brain MRI samples.

The confusion matrix shows the number of images correctly and incorrectly recognized by the model. For a detailed analysis and a complete understanding of the number of correct and misclassified cases for each individual TL model, see the Confusion matrices presented in Fig. 14. By observing the confusion matrix, the results obtained from the test set are good. The given models can be used to detect the presence of tumors in the human brain in real time.

The Xception architecture has proven to be superior to other architectures. The Xception confusion matrix (Fig. 14 (a)) showed that the developed model was able to detect 310 out of 311 brain tumor patients, 290 out of 291 normal patients as healthy. We can see that the best performing models (Xception and InceptionResNetV2) have very few false negative (FN) counts (i.e. 0 and 1), which helps

to increase the sensitivity value. FN indicates that the models identify a patient with a brain tumor as healthy, whereas the patient has a tumor. Second, the models also show very few False Positives (i.e. 0 and 1) misidentified as tumor patients, which ultimately contributes to higher specificity and precision values.

2) EXPERIMENTATION 2: CLASSIFICATION ACCURACY ON MRI-SMALL DATASET

To check robustness, we tested our proposed TL models on the MRI-small dataset. Detailed classification results from all TL models are compared in terms of various metrics and summarized in Table 11. It can be seen that the Xception model also performed well on the second dataset, achieving the highest performance with 96.55% sensitivity, 87.88% specificity, 91.80% F1-score and 91.94% accuracy. It was also noted that the obtained sensitivity and NPV for the tumor class is significantly higher (i.e. 97% and 96.67%) for the Xception model. NasNet Large shows 2nd best performance in terms of accuracy of 91.74%. It is worth noting that InceptionResNetV2 showed the best value for specificity and precision of 96.97% and 96.00%, but achieved a lower value for sensitivity.

The class-wise performance of the models is presented in Table 12. This shows that Xception achieved the best results

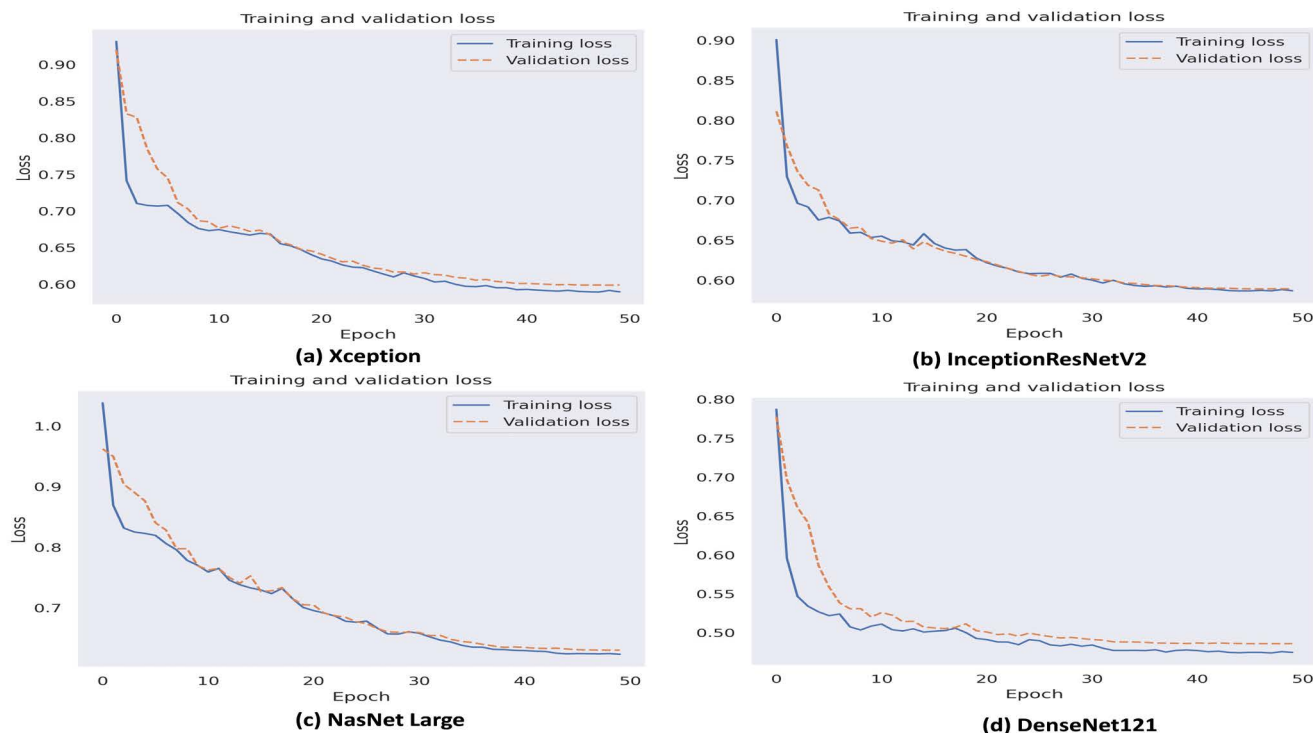


FIGURE 13. Epochs versus loss graph for all the models.

TABLE 11. Classification results comparison of all TL models on MRI-small dataset.

Model	ACC	SEN	PRE	SPE	F1-Score	NPV	FPR	FNR
Xception	91.94	96.55	87.50	87.88	91.80	96.67	0.121	0.034
NasNet Large	91.74	93.10	90.00	90.91	91.53	93.76	0.090	0.690
DenseNet121	90.32	89.66	89.66	90.91	89.66	90.91	0.090	0.103
InceptionResNetV2	90.32	82.76	96.00	96.97	88.89	86.49	0.030	0.172

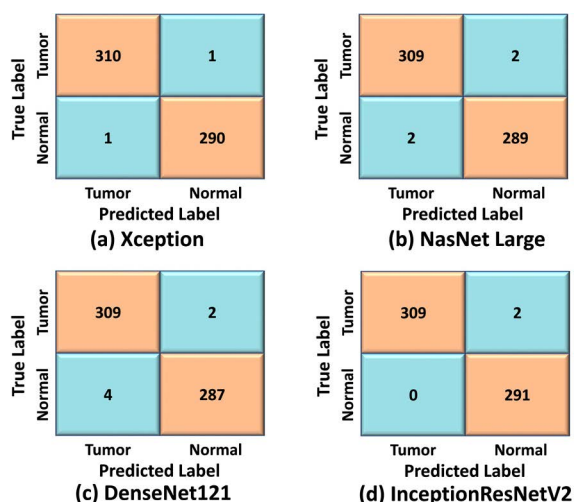


FIGURE 14. Confusion matrices obtained for all the models.

in the “tumor” category when calculating the sensitivity, and the precision of the “normal” category also reached 96.6%.

TABLE 12. Class-wise precision, recall and F1 score for all the models.

Model	Class	PRE	SEN	F1-score
Xception	Tumor	87.5	96.5	91.8
	Normal	96.6	87.8	92.0
NasNet Large	Tumor	90.0	93.1	91.5
	Normal	93.7	90.9	92.3
DenseNet121	Tumor	89.6	89.6	89.6
	Normal	90.9	90.9	90.9
InceptionResNetV2	Tumor	96.0	82.7	88.9
	Normal	86.4	96.9	91.3

Therefore, the F1-score of the Xception model was the best among all models. It can be seen that the NasNet Large model gives very good results in the classification of brain tumors. In addition, the results obtained using the Xception model are close to the results of NasNet Large, which is quite good. InceptionResNetV2 achieved the best results with the “normal” class when calculating sensitivity, and it measures 96% precision for the “tumor” class.

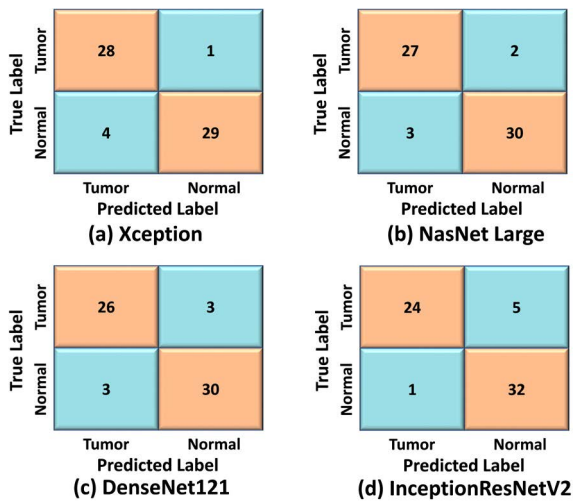


FIGURE 15. Confusion matrices obtained for all the models.

The Xception architecture has proven to be superior to other architectures. The Xception confusion matrix (Fig. 15 (a)) showed that the developed model was able to detect 28 out of 29 brain tumor patients, 29 out of 33 normal patients as healthy.

We can see that the best performing model (Xception) have very few false negative (FN) counts (i.e. 0 and 1), which helps to increase the sensitivity value. FN indicates that the models identify a patient with a brain tumor as healthy, whereas the patient has a tumor. Second, the model also shows very few false positives misidentified as tumor patients, which ultimately contributes to higher specificity and precision values.

E. DISCUSSION

The latest developments in medical imaging tools have facilitated health workers. Medical informatics research has the best options make good use of these exponentially growing volumes of data. Early detection options are essential for effective treatment of brain tumors. In this paper, we propose an enhanced deep learning model by comprehensively evaluating the effectiveness of four most effective CNN models (Xception, NasNet Large, DenseNet121, InceptionResNetV2) for brain tumor classification from MRI images. Extensive experiments were performed on two different MRI datasets (MRI-small and MRI-large) to determine the best performing model for automated brain tumor detection by considering several factors including three different optimizers. From the experimental results, it can be concluded that the deep transfer learning models Xception and InceptionResNetV2 models are very effective in classifying brain MRI images, and the classification accuracy is close. A detailed comparative analysis of all methods demonstrates the superiority of the Xception model. It can be seen that the Xception model achieved the highest performance with 99.68% precision, 99.66% specificity, 99.68% F1-score, 99.67% accuracy on MRI large dataset.

TABLE 13. Comparative analysis of the proposed system with state-of-the-art models.

Authors	Techniques used	Accuracy
Cheng et al. [36]	KNN & SVM	91.28%
Paul et al. [12]	CNN Model	91.43%
Abiwinanda et al. [27]	CNN Model	84.19%
Afshar et al. [37]	CapsNet	90.89%
Badža and Barjaktarović [19]	Four-layer CNN	97.39%
Rehman et al. [21]	Fine-tune VGG16	98.69%
M. Soltaninejad et al. [38]	Random Forest Classifier	86.00%
Chelghoum et al. [18]	Fine tuning	98.71%
Deepak and Ameer [10]	GoogleNet + transfer learning	97.10%
Mehrotra, R et al. [39]	AlexNet + transfer learning	99.04%
Kang, J et al. [40]	DenseNet-121 + ResNeXt-101	98.83%
	+ MnasNet	98.04%
	DenseNet-169	
Proposed Model	Xception (MRI-Large)	99.67%
	Xception (MRI-small)	91.94%

Similarly, Xception model also performed well on the second dataset (MRI-small dataset), achieving the highest performance with 96.55% sensitivity, 87.88% specificity, 91.80% F1-score and 91.94% accuracy.

F. COMPARISON WITH THE STATE-OF-THE-ART METHODS

This article presents a framework for selecting a pre-trained model that uses TL to classify brain tumors. The results achieved by the TL model are compared with the recently proposed method for automated brain tumor diagnosis using the same MRI dataset in Table 13. The results are compared based on models developed for brain tumor detection. It can be clearly seen from the table below that our proposed model based on the Xception architecture is significantly better than other state-of-the-art models in terms of evaluation indicators (such as accuracy).

V. CONCLUSION

In this study, we used transfer learning to develop a CNN model for automatic brain tumor diagnosis using MR images. Transfer learning uses weights from networks previously trained on millions of data. The proposed study implements four different transfer learning models with different optimizers (ADAM, SGD, RMSprop), and extensive experiments were performed on the two datasets with the largest number of MR images currently available. For these four models, the features are extracted using transfer learning, and three dense layers along with the softmax layer are used for classification purposes. The proposed deep TL models shows fast learning by using the Adam optimizer, and the dropout method avoids the problem of overfitting. The various proposed models were compared according to accuracy, recall, precision, and F1-score. After extensive experimentation, it is clear that deep transfer learning with the Xception model gave the best results among all TL models used in this study. On the benchmark datasets, e.g. MRI-large and MRI-small, For the Xception model, the

sensitivity scores were found to be 99.68% and 96.55%, respectively; precision scores were 99.68% and 87.50%, respectively; NPV scores were 99.65% and 96.67%, respectively; F-1 scores were recorded as 99.68% and 91.80%, respectively; accuracy was 99.67% and 91.94%, respectively. Our proposed model outperforms existing models with an accuracy of 99.67%. This demonstrates the effectiveness of our proposed method and the potential of using deep learning to quickly diagnose brain tumors through MRI. In future work, the performance of the system can still be improved by using larger data sets and using other deep learning techniques (such as GAN).

REFERENCES

- [1] D. N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: A summary," *Acta Neuropathol.*, vol. 131, no. 6, pp. 803–820, Jun. 2016.
- [2] M. I. Sharif, J. P. Li, M. A. Khan, and M. A. Saleem, "Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images," *Pattern Recognit. Lett.*, vol. 129, pp. 181–189, Jan. 2020.
- [3] J. J. Graber, C. S. Cobbs, and J. J. Olson, "Congress of neurological surgeons systematic review and evidence-based guidelines on the use of stereotactic radiosurgery in the treatment of adults with metastatic brain tumors," *Neurosurgery*, vol. 84, no. 3, pp. E168–E170, 2019.
- [4] A. Rehman, M. A. Khan, T. Saba, Z. Mehmood, U. Tariq, and N. Ayesha, "Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture," *Microsc. Res. Technique*, vol. 84, no. 1, pp. 133–149, 2021.
- [5] M. A. Khan, M. Sharif, T. Akram, S. A. C. Bukhari, and R. S. Nayak, "Developed Newton-raphson based deep features selection framework for skin lesion recognition," *Pattern Recognit. Lett.*, vol. 129, pp. 293–303, Jan. 2020.
- [6] M. A. Khan, M. S. Sarfraz, M. Alhaisoni, A. A. Albeshar, S. Wang, and I. Ashraf, "StomachNet: Optimal deep learning features fusion for stomach abnormalities classification," *IEEE Access*, vol. 8, pp. 197969–197981, 2020.
- [7] P. M. S. Raja and A. V. Rani, "Brain tumor classification using a hybrid deep autoencoder with Bayesian fuzzy clustering-based segmentation approach," *Biocybernetics Biomed. Eng.*, vol. 40, no. 1, pp. 440–453, Jan. 2020.
- [8] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, and J. Lu, "Brain tumor classification for MR images using transfer learning and fine-tuning," *Computerized Med. Imag. Graph.*, vol. 75, pp. 34–46, Jul. 2019.
- [9] K. Muhammad, S. Khan, J. D. Ser, and V. H. C. D. Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 507–522, Feb. 2021.
- [10] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput. Biol. Med.*, vol. 111, Aug. 2019, Art. no. 103345.
- [11] M. Toğaçar, B. Ergen, and Z. Cömert, "BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model," *Med. Hypotheses*, vol. 134, Jan. 2020, Art. no. 109531.
- [12] J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri, "Deep learning for brain tumor classification," *Proc. SPIE*, vol. 10137, Mar. 2017, Art. no. 1013710.
- [13] A. Çınar and M. Yildirim, "Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture," *Med. Hypotheses*, vol. 139, Jun. 2020, Art. no. 109684.
- [14] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, "Multi-grade brain tumor classification using deep CNN with extensive data augmentation," *J. Comput. Sci.*, vol. 30, pp. 174–182, Jan. 2019.
- [15] S. Kumar, A. Sharma, and T. Tsunoda, "Brain wave classification using long short-term memory network based OPTICAL predictor," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, Dec. 2019.
- [16] N. V. Shree and T. N. R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network," *Brain Inform.*, vol. 5, no. 1, pp. 23–30, 2018.
- [17] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, and M. Shoaib, "A deep learning model based on concatenation approach for the diagnosis of brain tumor," *IEEE Access*, vol. 8, pp. 55135–55144, 2020.
- [18] R. Chelghoum, A. Ikhlef, A. Hameurlaine, and S. Jacquir, "Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images," in *Proc. 16th IFIP WG 12.5 Int. Conf. Artif. Intell. Appl. Innov. (AIAI)*, Neos Marmaras, Greece, vol. 583, Jun. 2020, pp. 189–200.
- [19] M. M. Badža and M. Č. Barjaktarović, "Classification of brain tumors from MRI images using a convolutional neural network," *Appl. Sci.*, vol. 10, no. 6, p. 1999, Mar. 2020.
- [20] F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthcare*, vol. 9, no. 2, p. 153, Feb. 2021.
- [21] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Syst., Signal Process.*, vol. 39, no. 2, pp. 757–775, Feb. 2020.
- [22] M. Talo, U. B. Baloglu, Ö. Yıldıırım, and U. R. Acharya, "Application of deep transfer learning for automated brain abnormality classification using MR images," *Cogn. Syst. Res.*, vol. 54, pp. 176–188, May 2019.
- [23] H. M. Ahmed, B. A. B. Youssef, A. S. Elkorany, A. A. Saleeb, and F. A. El-Samie, "Hybrid gray wolf optimizer-artificial neural network classification approach for magnetic resonance brain images," *Appl. Opt.*, vol. 57, no. 7, p. B25, 2018.
- [24] P. Saxena, A. Maheshwari, and S. Maheshwari, "Predictive modeling of brain tumor: A deep learning approach," in *Innovations in Computational Intelligence and Computer Vision*. Singapore: Springer, 2021, pp. 275–285.
- [25] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, "Multi-class brain tumor classification using residual network and global average pooling," *Multimedia Tools Appl.*, vol. 80, no. 9, pp. 13429–13438, Apr. 2021.
- [26] S. Khawaldeh, U. Pervaiz, A. Rafiq, and R. Alkhalwaleh, "Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks," *Appl. Sci.*, vol. 8, no. 1, p. 27, Dec. 2017.
- [27] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in *Proc. World Congr. Med. Phys. Biomed. Eng.* Singapore: Springer, 2018, pp. 183–189.
- [28] F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthcare*, vol. 9, no. 2, p. 153, Feb. 2021.
- [29] A. Hamada. (2020). *Br35H: Brain Tumor Detection 2020, Version 5*. [Online]. Available: <https://www.kaggle.com/ahmedhamada0/brain-tumor-detection>
- [30] N. Chakrabarty. (2019). *Brain MRI Images Dataset for Brain Tumor Detection*, Kaggle. [Online]. Available: <https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection>
- [31] A. Rosebrock, *Finding Extreme Points in Contours With Open CV*. [Online]. Available: <https://www.pyimagesearch.com/2016/04/11/finding-extreme-points-in-contours-with-opencv/>
- [32] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.
- [33] Q. Zhang, Y. N. Wu, and S.-C. Zhu, "Interpretable convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8827–8836.
- [34] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
- [35] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, San Francisco, CA, USA, 2017, pp. 4278–4284.
- [36] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0140381.

- [37] P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1368–1372.
- [38] M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T. L. Jones, T. R. Barrick, F. A. Howe, and X. Ye, "Supervised learning based multimodal MRI brain tumour segmentation using texture features from supervoxels," *Comput. Methods Programs Biomed.*, vol. 157, pp. 69–84, Apr. 2018.
- [39] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, "A transfer learning approach for AI-based classification of brain tumors," *Mach. Learn. with Appl.*, vol. 2, Dec. 2020, Art. no. 100003.
- [40] J. Kang, Z. Ullah, and J. Gwak, "MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers," *Sensors*, vol. 21, no. 6, p. 2222, Mar. 2021.
- [41] J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "A distinctive approach in brain tumor detection and classification using MRI," *Pattern Recognit. Lett.*, vol. 139, pp. 118–127, Nov. 2020.
- [42] Ö. Polat and C. Güngen, "Classification of brain tumors from MR images using deep transfer learning," *J. Supercomputing*, vol. 77, no. 7, pp. 1–17, 2021.
- [43] H. Mohsen, E.-S. El-Dahshan, E.-S. El-Horbarty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Future Comput. Informat. J.*, vol. 3, pp. 68–71, Jun. 2018.
- [44] S. Kumar, C. Dabas, and S. Godara, "Classification of brain MRI tumor images: A hybrid approach," *Proc. Comput. Sci.*, vol. 122, pp. 510–517, Dec. 2017.
- [45] F. Provost and R. Kohavi, "Glossary of terms," *J. Mach. Learn.*, vol. 30, nos. 2–3, pp. 271–274, 1998.
- [46] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: An overview," 2020, *arXiv:2008.05756*.



SOHAIB ASIF received the bachelor's degree in computer science from COMSATS University, Pakistan, in 2017, and the master's degree from Xi'an Jiaotong University, China, in 2020. He is currently pursuing the Ph.D. degree with Central South University, Changsha, China. He has authored one book chapter and more than four scientific research papers. His research interests include medical image analysis, image classification, deep learning, and machine learning.

He received three Best Presentation Awards in IEEE ICC 2020 and 2021 and IEEE ICAIBD 2020.



WENHUI YI received the bachelor's, master's, and Ph.D. degrees from Xi'an Jiaotong University, in 1996, 1999, and 2004, respectively. He is an Associate Professor with the Faculty of Information and Electronics Engineering, School of Electronics Science and Engineering, Xi'an Jiaotong University. He has been working on developing biosensors and the intelligent diagnosis and treatment system toward cancers, including diagnosis of cancers by classifying pathological

images utilizing deep convolutional neural networks, in the past five years. He was a Postdoctoral Fellow with the University of Akron, from 2007 to 2008. He was a Visiting Scientist at MIT, from 2011 to 2013.



QURRAT UL AIN received the master's degree from Arid Agriculture University, Pakistan, in 2019. She is currently pursuing the Ph.D. degree with the School of Public Health, Central South University, Changsha, China. Her research interests include public health, intelligent systems for healthcare applications, and medical statistics.



JIN HOU received the master's degree from Xi'an Jiaotong University, in 2000, and the Ph.D. degree from Fourth Military Medical University, in 2006. She joined as an Assistant Professor with the Xi'an Medical College, in 2000, and was promoted to Associate Professor, in 2008, and to Full Professor, in 2016. She was a Visiting Scholar at the Whitehead Institute, MIT, in 2013, and at the University of Houston, in 2017. She is working on screening compounds for antitumor

drugs. Her research interests include muscarinic receptors and intracellular signal transduction.



TAO YI is currently pursuing the undergraduate degree with the Faculty of Electronics and Information Engineering, School of Computer Science and Technology, Xi'an Jiaotong University. He joined this research through the Bud Program for Talented Undergraduates of Xi'an Jiaotong University. His research interests include the domain of biomedical engineering, computer vision, and deep learning.



JINHAI SI received the bachelor's and master's degrees from Liaoning University, in 1982 and 1985, respectively, and the Ph.D. degree from the Harbin Institute of Technology, in 1994. He was a Postdoctoral Fellow at the Institute of Physics, Chinese Academy of Sciences, from 1994 to 1996, and at the Tokyo Institute of Technology, from 1988 to 1990. He was a Lecturer and an Associate Professor with Liaoning University, from 1985 to 1994, and an Associate

Professor with the Institute of Physics, Chinese Academy of Sciences, from 1996 to 1999, and a Chief Researcher and the Team Leader of Japan Science and Technology Agency, from 1997 to 2005. He has been a Full Professor with Xi'an Jiaotong University, since 2005. His research interests include ultrafast optical spectroscopies and techniques, femtosecond laser-based micro-nano machining, and transient imaging and ultrafast measurements using femtosecond optical Kerr gate.

...